Estimation of linear models with inequality restrictions

Moors, J.J.A.; van Houwelingen, J.C.

Publication date:
1987

TILBURG UNIVERSITY
UNIVERSITEIT BRABANT
POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

[redacted]

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

[redacted]

ESTIMATION OF LINEAR MODELS WITH INEQUALITY RESTRICTIONS

J.J.A. Moors
J.C. van Houwelingen

FEW 291

ESTIMATION OF LINEAR MODELS WITH

INEQUALITY RESTRICTIONS

J.J.A. Moors* & J.C. van Houwelingen**

* Tilburg University, P.O. Box 90153,5000 LE Tilburg, The Netherlands
** Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands

ESTIMATION OF LINEAR MODELS WITH INEQUALITY RESTRICTIONS

J.J.A. Moors & J.C. van Houwelingen

Summary Inequality constrained regression involves the notion of a truncated parameter space, which was studied extensively in MOORS (1985). His general results are extended here and applied to linear models. Using the invariance principle, for every observation x a set V_x is defined with the property that estimates in V_x necessarily correspond to inadmissible estimators. One of the practical conclusions is that the usual estimators in inequality constrained regression are inadmissible.

1. Introduction

An appetizer to start with. Consider the simple regression model $y = \vartheta_1 + \vartheta_2 x + \varepsilon$ with $\varepsilon \sim N(0,1)$, where it is known a priori that $|\vartheta_2| \le 0.5$. Only (the plausible class of) symmetric estimators $\hat{\vartheta}$ will be taken into consideration, defined by the property that $\hat{\vartheta}$ changes sign if all observations y_i change sign. Five observations produce the sums

$$\Sigma\, y_i = 1 \qquad \Sigma\, x_i = 10$$
$$\Sigma\, y_i x_i = 1.6 \qquad \Sigma\, x_i^2 = 21$$

What is wrong now with the estimate (1, -0.4) for $(\vartheta_1,\vartheta_2)$? It will be shown here that this estimate necessarily corresponds with an inadmissible estimator, which means that a better (symmetric) estimator exists. The important point is of course, that this conclusion can be drawn without regard to be behaviour of the estimator for other data. Note that the familiar least squares estimator is among these inadmissible estimators!

Linear regression problems where the coefficients have to satisfy a priori given inequality restrictions occur rather frequently and pose interesting but difficult problems. The practical solution to this kind of estimation problem usually consists of two steps:

a) apply some classical estimation method;

b) (i) if the resulting estimates satisfy the inequality constraints, they remain unchanged;

(ii) if not, the estimates are adapted in such a way that the new estimates do meet the constraints, while being 'close' to the original ones.

Although this general estimation procedure is common practice, statistical properties of the resulting estimators are as yet unknown. Particularly (ii) - which is in fact a projection on the feasible set of the coefficients - makes it very hard to determine whether the estimators obtained are unbiased, admissible, and so on.

The inequality constraints on the regression parameters lead to a so-called truncated parameter space, which means that part of the theoretically possible parameter values is excluded beforehand. Truncated parameter spaces in general were studied in MOORS (1985). In his Theorems 2.2 and 2.3 it was shown that under rather mild conditions no unbiased estimators exist for parameters in truncated spaces. In Chapter 3, invariant estimation problems were studied and fairly generally the inadmissibility was proved of estimators taking values on, or even near, the boundary of the parameter space. An earlier version of this result is MOORS (1981).

To make the paper reasonably self-contained these previous findings will be summarized in the remainder of this section. Since the underlying assumptions appear to be unnecessarily restrictive, it is shown in Section 2 which assumptions can safely be dropped. In the next two sections this general theory is applied to inequality constrained regression; the standard linear model with $\varepsilon \sim N_n(0,\sigma^2 I)$ is discussed: the case of known σ^2 in Section 3, unknown σ^2 in Section 4. Two types of truncated parameter spaces are considered: blocks and ellipsoids. The final Section 5 discusses the results and looks ahead at future work.

As the usual definition of a decision problem - see FERGUSON (1967) or LEHMANN (1983) for example - is felt to be too general in estimation theory, here the starting point is the following definition. The convex closure (closed convex hull) of a set S will be denoted by C{S}.

Definition 1 An estimation problem (Θ, L, X) consists of the following elements:

(i) a given parameter space $\Theta \subset \mathbf{R}_k$;
(ii) a given function $h: \Theta \to \mathbf{R}_m$;
(iii) the action space $A = C\{h(\Theta)\} \subset \mathbf{R}_m$;
(iv) a given loss function $L: \Theta \times A \to \mathbf{R}$;
(v) a random observable X with some probability distribution P_ϑ, $\vartheta \in \Theta$.

Further, L is measurable in the pair (ϑ, a), allowing a representation

$$(1.1) \quad L(\vartheta, a) = L^*(\vartheta, h(\vartheta) - a)$$

and $h(\Theta)$ contains an open set in $\mathbf{R}_m$.

In the special case of quadratic loss function

$$(1.2) \quad L(\vartheta, a) = |h(\vartheta) - a|^2$$

the estimation problem is called quadratic. □

Note the sensible assumption that the loss only depends on a through $h(\vartheta) - a$, where $h(\vartheta)$ is the estimand, and the choice of A as the convex closure of $h(\Theta)$.

Invariant estimation problems are standardly characterized by the existence of a group G of functions g with corresponding groups $\bar{G}$ and $\tilde{G}$ such that for all $g \in G$ a (unique) $\bar{g} \in \bar{G}$ and $\tilde{g} \in \tilde{G}$ exist with the properties

$$(1.3) \quad P_{\bar{g}(\vartheta)}\{X \in B\} = P_\vartheta\{X \in g^{-1}(B)\}$$

for all (Borel) sets B, and

$$(1.4) \quad L(\bar{g}(\vartheta), \tilde{g}(a)) = L(\vartheta, a)$$

for any $a \in A$ and all $\vartheta \in \Theta$. An estimator d is called invariant if in addition

(1.5) $dg(x) = \tilde{g}d(x)$

holds for all realisations x of X and any $g \in G$. There are good reasons to use invariant estimators for invariant problems.

Only <u>dominated</u> classes $\{P_\vartheta : \vartheta \in \Theta\}$ will be considered here, meaning that a σ-finite measure μ exists with corresponding densities $f(x|\vartheta)$. Function g is called <u>measure preserving</u> if

(1.6) $\mu\{g^{-1}(B)\} = \mu\{B\}$

holds for all B. A function $\tilde{g} \in \tilde{G}$ will be called <u>linear</u>, if $\tilde{g}(\sum_i c_i x_i) = \sum c_i \tilde{g}_i(x_i)$ holds for all positive scalars c_i with $\sum c_i = 1$. These notions are used to define a specific type of invariant estimation problems.

<u>Definition 2</u> Let estimation problem (Θ, L, X) be invariant under a finite group G (with induced group $\bar{G}$ and $\tilde{G}$); let $\{P_\vartheta : \vartheta \in \Theta\}$ be dominated. If

(i) all $g \in G$ are measure preserving,
(ii) all $\tilde{g} \in \tilde{G}$ are linear, and
(iii) $\tilde{G}$ is commutative

then the problem is called <u>linearly invariant</u> or <u>linvariant</u> for short. Invariant estimators for linvariant problems are called <u>linvariant</u> as well. □

For a linvariant problem the following properties can be proved for any $g \in G$:

(1.7) $h\bar{g} = \tilde{g}h$

(1.8) $f(x|\bar{g}(\vartheta)) = f(g^{-1}(x)|\vartheta)$

a.e. (with respect to μ).

For notational convenience the functions g now are supplied with a suffix i, so that $G = \{g_i : i \in I\}$; let Σ denote summation over all $i \in I$. For any fixed x a function $h_x : \Theta \to A$ and a set $A_x \subset \mathbb{R}_m$ are defined as follows:

$$(1.9)\qquad h_x(\vartheta) := \begin{cases} \dfrac{\Sigma f(x|\bar{g}_i(\vartheta)).\tilde{g}_i h(\vartheta)}{\Sigma f(x|\bar{g}_i(\vartheta))} & \text{for } \Sigma f(x|\bar{g}_i(\vartheta)) > 0 \\ h(\vartheta) & \text{for } \Sigma f(x|\bar{g}_i(\vartheta)) = 0 \end{cases}$$

$$(1.10)\qquad A_x := C\{h_x(\Theta)\}$$

They can be shown to satisfy

$$(1.11)\qquad h_{g(x)}\bar{g} = \tilde{g}h_x$$

$$(1.12)\qquad A_{g(x)} = \tilde{g}(A_x)$$

Besides, h_x is a contraction, meaning that $|h_x(\vartheta)| \le |h(\vartheta)|$ holds for all $\vartheta \in \Theta$; $A_x \subset A$ follows at once.

Now the main theorem can be proved.

Theorem 1 Consider linvariant quadratic estimation problem (Θ, L, X) and let D_L denote the class of linvariant estimators. Assume that for $d \in D_L$ a $\vartheta \in \Theta$ exists for which $\{x : d(x) \not\subset A_x\}$ has positive probability. Then d is strictly dominated by $d_0 \in D_L$, where $d_0(x)$ is defined for all x as the projection of $d(x)$ on A_x. □

The theorem states that any linvariant estimator taking a value outside A_x for some observation x can be improved by projecting this value on A_x. Note that this holds true, whatever the behaviour of the estimator for other observations. Hence, for any x the action space A can in fact be reduced to A_x. Of course, this conclusion is of interest only if A_x is a strict subset of A - at least for some x. In that case all estimators taking values on (or: close enough to) the boundary of A are inadmissible. This applies, for example, to all regression estimators obtained by the general procedure described in the beginning of this section. Up to now, the most interesting applications have been derived for truncated parameter spaces. If for two estimation problems (Θ_0, L, X) and (Θ, L, X), Θ_0 is a strict subset of Θ, the parameter space Θ_0 is called truncated (with respect to Θ).

Another way to formulate Theorem 1 is by means of the following notion. A random set $V_X \subset \mathbb{R}_m$ is a function of X that assigns a set V_x to any realisation $x \in X$.

Definition 3 Let D_0 be a class of estimators for the estimation problem (Θ, L, X). Random set $V_X \subset \mathbb{R}_m$ is called inadmissible under D_0, if any estimator $d \in D_0$ with the property

$$P_\vartheta\{d(X) \in V_X\} > 0$$

for some $\vartheta \in \Theta$ is strictly dominated by some $d_0 \in D_0$. □

Then the conclusion of Theorem 1 comes down to the statement that $A-A_X$ is inadmissible under D_L.

2. Generalisations

The assumptions of the previous section are unnecessarily restrictive. It will be shown here that Theorem 1 remains valid if the conditions (i) and (iii) in Definition 2 are dropped.

As to (i), μ dominates $\{P_\vartheta : \vartheta \in \Theta\}$. Define the function μ^* on the (Borel) sets B of the sample space by

$$\mu^*(B) := \frac{1}{m} \Sigma \mu\{g_i(B)\}$$

where m is the number of elements of G. It is easily checked that μ^* is a σ-finite measure again; further μ^* dominates $\{P_\vartheta : \vartheta \in \Theta\}$. Finally, since G is a group, $\{g_i g^{-1} : i \in I\} = G$, so that

$$\mu^*\{g^{-1}(B)\} = \frac{1}{m}\,\Sigma\,\mu\{g_i g^{-1}(B)\}$$

$$= \frac{1}{m}\,\Sigma\,\mu\{g_i(B)\} = \mu^*\{B\}$$

holds for all B. Hence the measure μ^* is preserved by all $g \in G$. So, condition (i) can always be satisfied by replacing μ by μ^* eventually.

Condition (iii) of Definition 2 was used in MOORS (1985) only to prove (1.11). The following alternative proof avoids the use of (iii).

For $\Sigma f(x|\bar{g}_i(\theta)) > 0$, the following chain of equalities holds.

$$h_{g(x)}\bar{g}(\theta) = \frac{\Sigma f(g(x)|\bar{g}_i\bar{g}(\theta)).\tilde{g}_i h\bar{g}(\theta)}{\Sigma f(g(x)|\bar{g}_i\bar{g}(\theta))}$$

$$= \frac{\Sigma f(x|\bar{g}^{-1}\bar{g}_i\bar{g}(\theta)).h\bar{g}_i\bar{g}(\theta)}{\Sigma f(x|\bar{g}^{-1}\bar{g}_i\bar{g}(\theta))}$$

$$= \frac{\Sigma f(x|\bar{g}^{-1}\bar{g}_i\bar{g}(\theta)).h\overline{gg^{-1}}\bar{g}_i\bar{g}(\theta)}{\Sigma f(x|\bar{g}^{-1}\bar{g}_i\bar{g}(\theta))}$$

$$= \frac{\Sigma f(x|\bar{g}_i(\theta)).\tilde{g}h\bar{g}_i(\theta)}{\Sigma f(x)|\bar{g}_i(\theta))}$$

$$= \tilde{g}\,h_x(\theta).$$

Here, (1.7) is used repeatedly; further the successive equalities follow from (1.9), (1.8), $\overline{gg^{-1}}$ is the identity, $\{\bar{g}^{-1}\bar{g}_i\bar{g}: i \in I\} = \bar{G}$ and, finally, the linearity of $\tilde{g}$.

For $\Sigma f(x|\bar{g}_i(\theta)) = 0$, the equalities

$$h_{g(x)}\bar{g}(\theta) = h\bar{g}(\theta) = \tilde{g}h(\theta) = \tilde{g}h_x(\theta)$$

hold, completing the proof of (1.11).

Hence, Theorem 1 remains valid even if the assumptions (i) and (iii) in Definition 2 are dropped.

3. Truncated linear model with known variance

The foregoing theory, in particular Theorem 1, now will be applied to the general linear regression model with (known) restrictions on the regression parameters. So, consider the model

$$(3.1)\qquad W = V\beta + \varepsilon,\ \beta \in B$$

where W is the n-vector of observations of the endogeneous (random) variable, V is the $n \times k$-matrix of values of the k explanatory (deterministic) variables and β is the k-vector of regression coefficients. The vector ε of (unobservable) errors has the distribution $N_n(0,\sigma^2 I_n)$. The problem is to estimate the unknown parameters using the (scalar) quadratic loss function (1.2), where h is the identity.

In the standard linear model $B = \mathbb{R}_k$; here, the parameter space is assumed to be a strict subspace of $\mathbb{R}_k$. Two types of subspaces will be considered:

(i) B is the k-dimensional <u>block</u> $\{\beta \in \mathbb{R}_k : a \le A\beta \le b\}$, where the k-vectors a and b ($\ge a$) and the $k \times k$-matrix A are known; $\ge$ denotes inequality for each component. For the components of a and b the values $-\infty$ and $+\infty$, respectively, are allowed.

(ii) B is the <u>ellipsoid</u> $\{\beta \in \mathbb{R}_k : (\beta-a)^T A(\beta-a) \le b\}$, where again a is a k-vector and A a $k \times k$-matrix, but b now is a (non-negative) scalar, all of which are known.

First of all, the two resulting truncated estimation problems, will be rephrased in a simpler form, leading to parameter spaces which are centered around the origin, hence making the problems invariant.

In case (i), A may be taken non-singular without loss of generality. Using the convention $p + q = 0$ for $p = -\infty$ and $q = \infty$, and introducing the notations

$$Y := W - VA^{-1}(a+b)/2$$

$$X := VA^{-1}$$

$$\vartheta := A\beta - (a+b)/2$$

$$c := (b-a)/2$$

the problem of estimating β within a block may be written as

$$(3.2) \qquad Y = X\vartheta + \varepsilon, \qquad \vartheta \in \Theta := \{\vartheta \in \mathbb{R}_k : -c \le \vartheta \le c\}$$

In case (ii), A may be taken symmetric, so that $A = P^T DP$ with D diagonal and P orthogonal (c.f. GRAYBILL (1969), Th. 3.4.4). With the notations

$$Y := W - Va$$

$$X := VP^T$$

$$\vartheta := P(\beta - a)$$

the problem of estimating β within an ellipsoid may be written as

$$(3.3) \qquad Y = X\vartheta + \varepsilon, \qquad \vartheta \in \Theta := \{\vartheta \in \mathbb{R}_k : \vartheta^T D\vartheta \le b\}$$

Only the formulations (3.2) and (3.3) will be considered in the sequel. Note that these two estimation problems can be considered as equivalent to the original pair.

In the remainder of this section the situation that σ^2 is known will be treated; without loss of generality, $\sigma^2 = 1$ will be assumed.

Both the estimation problems (3.2) and (3.3) are linvariant with respect to the group $G = \{e,g\}$, where g is defined by $g(y) = -y$. The corresponding functions $\bar{g}$ and $\tilde{g}$ are given by $\bar{g}(\vartheta) = -\vartheta$ and $\tilde{g}(a) = -a$, respectively. Indeed, (1.3) and (1.4) are satisfied:

$$g(y) = -y \sim N_n(-X\vartheta,\ I_n) = N_n(X\bar{g}(\vartheta),\ I_n)$$

$$L(\bar{g}(\vartheta),\ \tilde{g}(a)) = |\bar{g}(\vartheta) - \tilde{g}(a)|^2 = |-\vartheta + a|^2 = L(\vartheta, a)$$

while $\bar{g}(\Theta) = \Theta$.

Hence, Theorem 1 is applicable. From the density

$$f(y|\vartheta) = (2\pi)^{-k/2} \exp[-\tfrac{1}{2}\ |y-X\vartheta|^2]$$

it follows

$$f(y|\bar{g}(\vartheta)) = (2\pi)^{-k/2} \exp[-\tfrac{1}{2}\ |y+X\vartheta|^2]$$

(3.4) $$f(y|\vartheta)/f(y|\bar{g}(\vartheta)) = \exp[2y^{\mathsf{T}}X\vartheta]$$

Now (1.9) gives $h_y(\vartheta) = [f(y|\vartheta)\vartheta + f(y|\bar{g}(\vartheta))\tilde{g}(\vartheta)]/[f(y|\vartheta) + f(y|\bar{g}(\vartheta))]$ or

(3.5) $$h_y(\vartheta) = \vartheta \tanh[y^{\mathsf{T}}X\vartheta].$$

Note that $y^{\mathsf{T}}X$, the statistic occuring in (3.5), is sufficient for ϑ.

Next, the space $h_y(\Theta)$ has to be found for Θ given in (3.2) and (3.3). The next lemma shows that it suffices to find the h_y-image of the boundary of Θ, reflecting the fact that h_y is a contraction. Let $Bd\{S\}$ denote the boundary of a bounded subspace S of $\mathbb{R}_k$.

<u>Lemma 1</u> Let $S \subset \mathbb{R}_k$ be bounded and symmetric with respect to the origin. Then h_y defined by (3.5) satisfies $h_y(S) \subset S$ and

(3.6) $$Bd\ \{h_y(S)\} = h_y(Bd\{S\})$$

<u>Proof</u> Since the case $y^{\mathsf{T}}X = 0$ is trivial, assume $y^{\mathsf{T}}X \neq 0$. Let L be any line through the origin, so $L = \{k\ \vartheta_0 : k \in \mathbb{R}\}$ for some $\vartheta_0 \neq 0$; consequently, $L \cap S = \{k\ \vartheta_0 \colon |k| \leq k_0\}$ for some $k_0 \geq 0$. Then $h_y(L) \subset L$ and $h_y(L \cap S) \subset L \cap S$, since h_y is a contraction. As this holds for all L, $h_y(S) \subset S$. Note that $h_y(L) = \{0\}$ if ϑ_0 satisfies $y^{\mathsf{T}}X\vartheta_0 = 0$.

Since h_y is an even function of ϑ and monotone in $|\vartheta|$, $h_y(L \cap S)$ is the segment of L with endpoints 0 and $h_y(\pm k_0\vartheta_0)$. From $\pm k_0\vartheta_0 \in \mathrm{Bd}\{S\}$, (3.6) follows. □

The lemma implies that to find $h_y(\Theta)$ for Θ in (3.2), only vectors $\vartheta \in \Theta$ have to be considered with the property $\vartheta_i = c_i$ for at least one $i = 1,2,\dots,k$. In (3.3), $h_y(\vartheta)$ with $\vartheta^T D\vartheta = b$ leads to the boundary of $h_y(\Theta)$.

Lemma 2 shows that $h_y(\Theta)$ is convex, whenever Θ is convex. For the two types of parameter spaces considered here, it follows that the reduced action space A_y equals $h_y(\Theta)$ throughout this section.

<u>Lemma 2</u> Let $S \subset \mathbf{R}_k$ be convex and symmetric with respect to the origin. Then $h_y(S)$ with h_y defined by (3.5) is convex as well. □

<u>Proof</u> The proof of Lemma 1 showed that any point on the boundary of S satisfies $h_y(s) = ks$ for some $s \in \mathrm{Bd}\{S\}$ and $0 \le k \le 1$. Consider two such points $h_y(s_i) = k_i s_i$ $(i = 1,2)$; then the definition

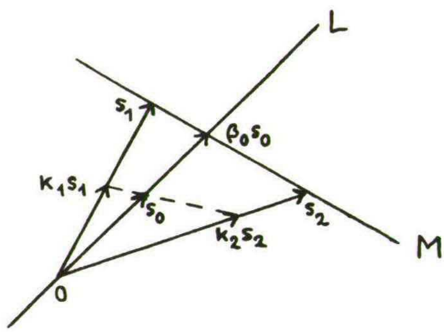
$$t_i := |y^T X s_i|, \quad i = 1,2$$

implies $k_i = \tanh t_i$. To prove the convexity of $h_y(S)$ it has to be shown that

$$s_0 := \alpha h_y(s_1) + (1-\alpha)h_y(s_2) \in h_y(S)$$

Let $\beta_0 s_0$ be the point of intersection of the lines $L := \{\beta s_0 : \beta \in \mathbf{R}\}$ and $M := \{\gamma(s_1 - s_2) + s_1 : \gamma \in \mathbf{R}\}$; compare the sketch. Then $\beta_0 s_0 \in S$ because of the convexity of S and $\beta_0 = [\alpha k_1 + (1-\alpha)k_2]^{-1}$. The proof is completed if it can be shown that

$$|h_y(\beta_0 s_0)| \ge |s_0|$$



holds, or equivalently,

$$\beta_0 \tanh\,[\beta_0(\alpha k_1 t_1+(1-\alpha)k_2 t_2)] \geq 1$$

or, again equivalently, using the inverse function F of tanh:

$$\alpha k_1 F(k_1) + (1-\alpha)k_2 F(k_2) \geq \frac{1}{\beta_0}F(\frac{1}{\beta_o}) = [\alpha k_1+(1-\alpha)k_1]F(\alpha k_1+(1-\alpha)k_1)$$

This last inequality comes down to the convexity for $|t| \leq 1$ of the function

$$G(t):= tF(t) = \frac{t}{2}\log\frac{1+t}{1-t}$$

which is easily checked. □

From now on, attention will be concentrated on the two-dimensional case. For k = 2, (3.5) can be rewritten as

$$h_y(\vartheta) = \begin{bmatrix}\vartheta_1\\ \vartheta_2\end{bmatrix}\tanh[a_1\vartheta_1+a_2\vartheta_2]$$

where $a:= (a_1,a_2)^T:= y^T X$. First, estimation problem (3.2) will be discussed, starting with the situation that $c_1 = \infty$.

In (3.2) with $c = (\infty,c_2)^T$, ϑ_1 is in fact unrestricted. Bd{Θ} consists of the lines $\vartheta_2 = \pm c_2$, so that boundary points of $h_y(\Theta)$ coincide with the set

$$(3.7)\qquad \left\{ \begin{bmatrix}k\\ c_2\end{bmatrix}\tanh[a_1 k+a_2 c_2]: k \in \mathbb{R} \right\}$$

Using the inverse function F of tanh, introduced above, and eliminating k, it can be shown that the boundary of $h_y(\Theta)$ is given by the curve

$$(3.8)\qquad \vartheta_1 = \frac{\vartheta_2}{a_1}\left[\frac{1}{2c_2}\log\frac{c_2+\vartheta_2}{c_2-\vartheta_2} - a_2\right]$$

(Since on this curve

$$\frac{\partial^2(a_1\theta_1)}{\partial\,\theta_2^2} = 2\left[\frac{c_2}{c_2^2 - \theta_2^2}\right]^2 > 0$$

holds, $h_y(\Theta)$ is convex indeed.)

Figure 1 shows the space A_y for a given value of c_2 and two values of a; compare MOORS (1985), Figure 3.23. Note that Θ is reduced substantially. For the numerical observations

$$a_1 = \Sigma\, x_{1i}y_i = 1 \qquad \Sigma\, x_{1i} = 5 \qquad \Sigma\, x_{2i} = 10$$

$$a_2 = \Sigma\, x_{2i}y_i = 1.6 \qquad \Sigma\, x_{1i}^2 = 5 \qquad \Sigma\, x_{2i}^2 = 21$$

with $n = 5$, mentioned in the introduction, the least squares estimate becomes $\hat{\theta} = (1, -0.4)$. Since this estimate is outside A_y, even the least squares estimator is inadmissible here.

Next, problem (3.2) will be studied with restrictions on both components of θ. The boundary of $h_y(\Theta)$ now further consists of the h_y-image of the lines $\theta_1 = \pm\, c_1$, namely the set

$$\left\{ \begin{bmatrix} c_1 \\ k \end{bmatrix} \tanh[a_1c_1+a_2k]\colon k \in \mathbb{R} \right\}$$

Similarly as before, this set is seem to be equivalent with the curve

$$(3.9) \qquad \theta_2 = \frac{\theta_1}{a_2}\left[\frac{1}{2c_1}\log\frac{c_1+\theta_1}{c_1-\theta_1} - a_1\right]$$

The space enclosed by (3.8) and (3.9) is $h_y(\Theta)$ and equals A_y by Lemma 2. The points of intersection of (3.8) and (3.9) are

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tanh\,[a_1c_1+a_2c_2], \quad \begin{bmatrix} c_1 \\ -c_2 \end{bmatrix} \tanh\,[a_1c_1-a_2c_2]$$

Figure 2 shows A_y for given c and a; the numerical values of c_2 and a are the same as in Figure 1.

To conclude this section, problem (3.3) is studied. For $k = 2$ write $D = \mathrm{diag}(d_1, d_2)$, so that

$$Bd\{\Theta\} = \{\theta \in \mathbb{R}_2 : d_1\theta_1^2 + d_2\theta_2^2 = b\}$$

In view of Lemma 1, the boundary points of $h_y(\Theta)$ constitute the set

$$\left\{ \begin{bmatrix} k \\ \sqrt{(b-d_1k^2)/d_2} \end{bmatrix} \tanh\,[a_1k + a_2\sqrt{(b-d_1k^2)/d_2}]: |k| \le b/d_1 \right\}$$

This set A_y is presented in Figure 3 for given b, d and a. To facilitate comparison, the values of b and d were chosen so that the resulting ellipse is the largest possible inside the parameter space of Figure 2. As to a, the same observations were taken as before.

4. Truncated linear model with unknown variance

The two estimation problems (3.2) and (3.3) will be discussed now in case σ^2 is unknown. The k-vector θ of unknown parameters must now be replaced by the (k+1)-vector

$$(4.1) \qquad \theta^* := (\theta^\top, \sigma^2)^\top$$

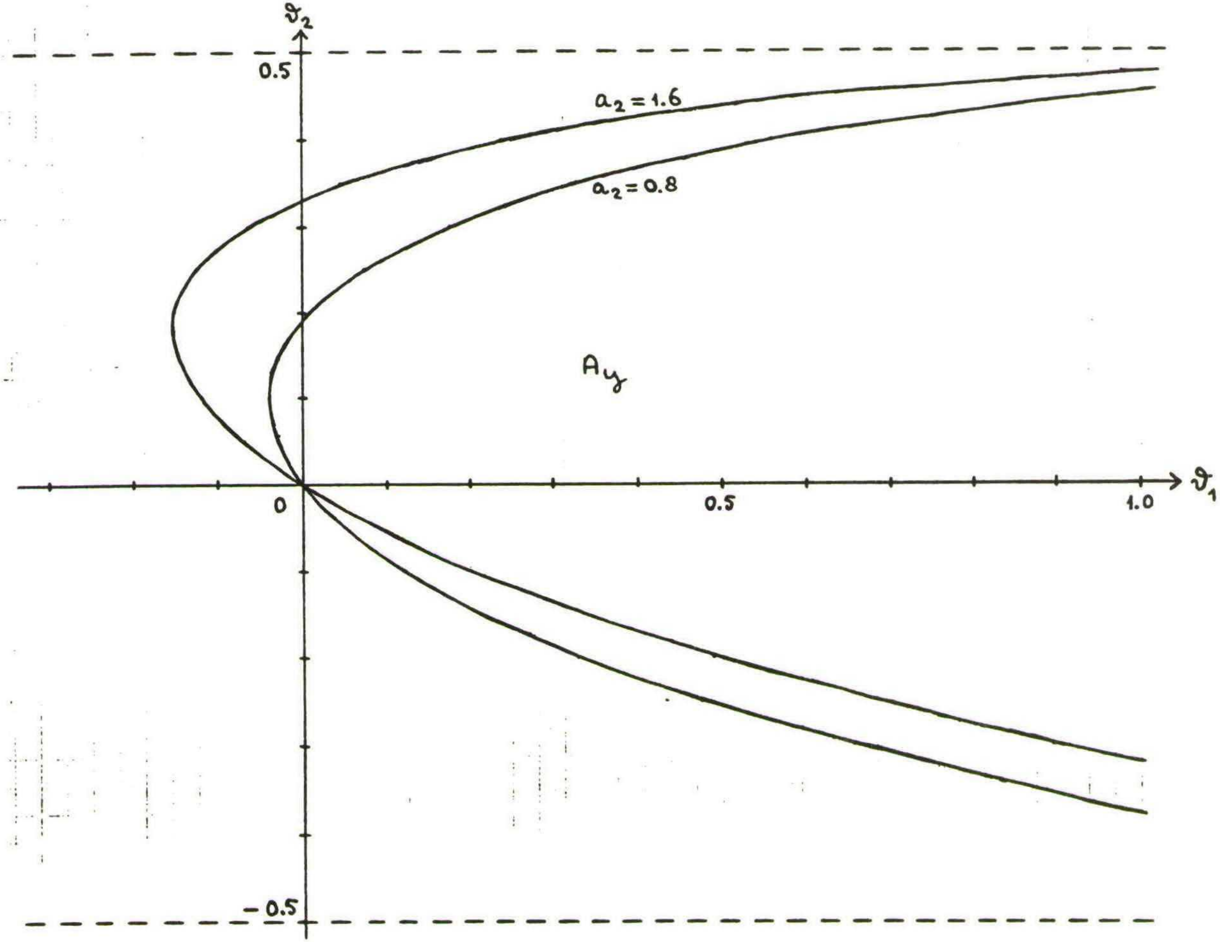
The parameter space Θ in (3.2) or (3.3) must be extended to

$$(4.2) \qquad \Theta^* := \Theta \times \mathbb{R}^+$$

if no a priori restrictions are put on σ^2.

Again the resulting problems are invariant under g defined by $g(y) = -y$; however, the corresponding $\bar{g}$ and $\tilde{g}$ are now both given by a $(k+1) \times (k+1)$-matrix:

$$(4.3) \qquad \bar{g} = \tilde{g} = \begin{bmatrix} -I_k & \emptyset \\ \emptyset & 1 \end{bmatrix}$$



<u>Figure 1</u> Reduced action space A_y for constraints $|\vartheta_2| \le 0.5$ ($a_1=1$)

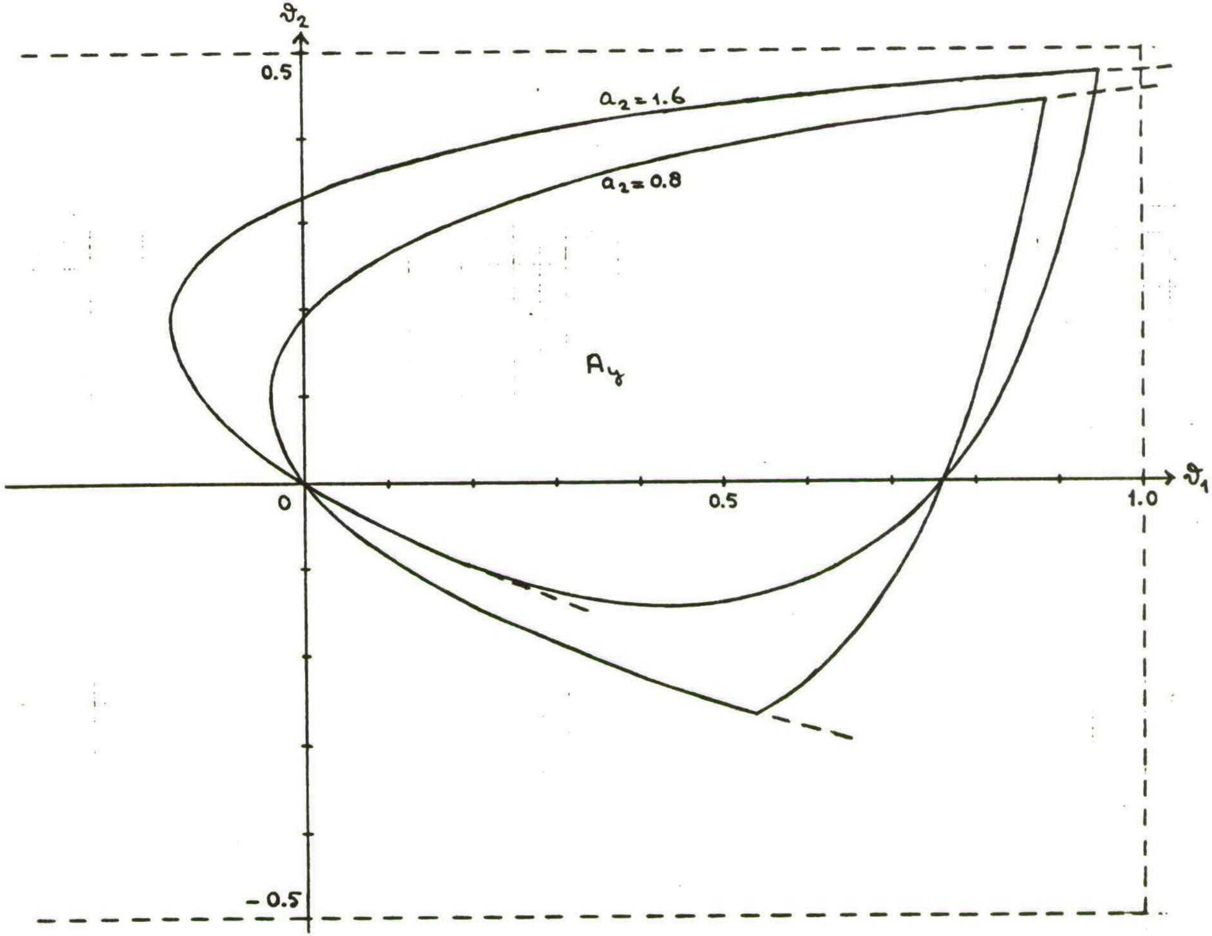
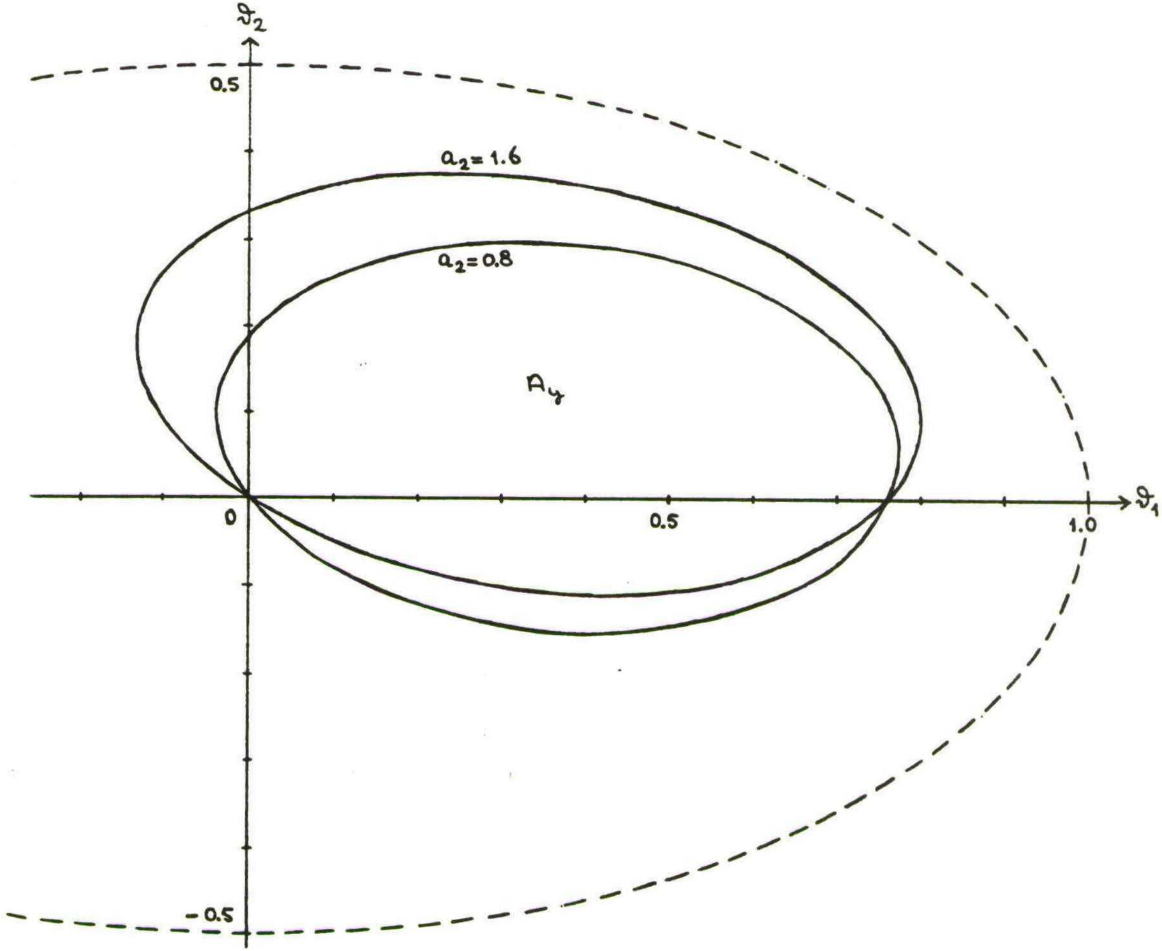


Figure 2 Reduced action space A_y for constraints $|\vartheta_1| \le 1$, $|\vartheta_2| \le 0.5$ ($a_1=1$)



<u>Figure 3</u> Reduced action space A_y for constraints $\vartheta_1^2 + 4\vartheta_2^2 \le 1$ $(a_1=1)$

Indeed, (1.4) is easily checked, while $y \sim P_{\vartheta^*} := N(X\vartheta, \sigma^2 I_n)$ implies

$$g(y) = -y \sim N(-X\vartheta, \sigma^2 I_n) = P_{g(\vartheta^*)}$$

so that (1.3) is satisfied too.

In stead of (3.4) comes the formula

$$f(y|\vartheta^*)/f(y|\bar{g}(\vartheta^*)) = \exp[2y^T X\vartheta/\sigma^2]$$

leading to

$$(4.4) \qquad h_y(\vartheta^*) = (\vartheta^T \tanh[y^T X\vartheta/\sigma^2],\ \sigma^2)^T$$

For the hyperplane $H_c := \{\vartheta^* \in \Theta^* : \sigma^2 = c\}$, $h_y(H_c) \subset H_c$ holds; comparison of (4.4) and (3.5) shows that pictures of $h_y(H_c)$ are similar to Figures 1-3. Further, $h_y(H_c) \subset h_y(H_{c'})$ for $c' \leq c$, since $|\vartheta^T \tanh[y^T X\vartheta/\sigma^2]|$ is decreasing in σ^2. Figure 4 gives an impression of $h_y(\Theta^*)$ based on the situation of Figure 3. Without additional constraints on σ^2, A_y is an elliptical cylinder; the truncation $c_1 \leq \sigma^2 \leq c_2$ reduces A_y to a skew cone.

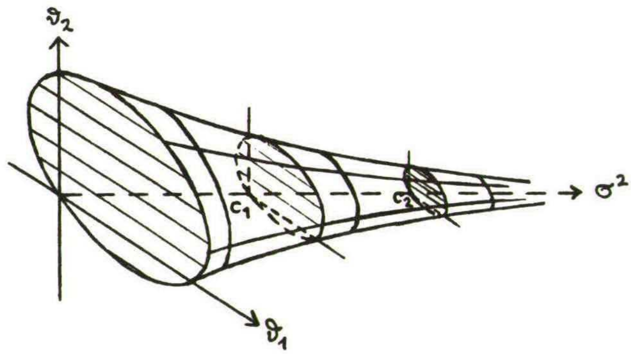


Figure 4 The space $h_y(\Theta^*)$ for unknown σ^2

Consider in more detail the case of one, constant, explanatory variable: $y = \vartheta + \varepsilon$, with $|\vartheta| \le m$. Then (4.4) reduces to

$$h_y \begin{bmatrix} \vartheta \\ \sigma^2 \end{bmatrix} = \begin{bmatrix} \vartheta \tanh[\vartheta \Sigma y_i/\sigma^2] \\ \sigma^2 \end{bmatrix}$$

so that

$$h_y(H_c) = \{(\vartheta, c)^T : 0 \le \vartheta \le m \tanh[m\ \Sigma y_i/\sigma^2]\}$$

Compare MOORS (1985), Example 4.7. Figure 5 shows A_y in a specific case.

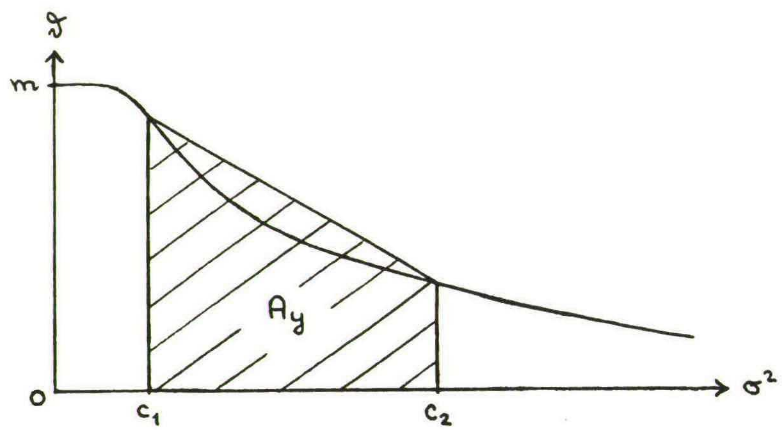


Figure 5 Reduced action space A_y for model $y = \vartheta + \varepsilon$

5. Discussion

A general approach to linvariant truncated estimation problems, showing the inadmissibility of estimators taking values near the boundary of the parameter space, was generalized by proving that two assumptions can be dropped.

Application of this general method to linear regression models with inequality restrictions on the parameters showed that the usual estimators devised for this type of problem are inadmissible. Since inequality constraint regression is an important issue (c.f. TOUTENBURG (1982), JUDGE et.al. (1982), Section 20.3 or HERING et.al. (1987)) these results are of interest.

The conclusions drawn are destructive in the sense that they show the serious drawbacks of existing estimators. However, Theorem 1 is essentially constructive: it indicates as well how to find a better estimator d_0. An open question is how good d_0 really is; an especially interesting question is its admissibility.

The only case considered here is h being the identity. However, Theorem 1 is applicable to other functions of θ as well. In regression two obvious applications are prediction (where $h(\theta) = \theta^T x_0$ for some given vector x_0) and the situation of Section 4 where no separate estimate of σ^2 is desired ($h(\theta^*) = \theta$).

Further, only symmetry with respect to the origin was considered, leading to the very simple group $G = \{e,g\}$ with $g(y) = -y$. Note that invariance occurs as well, if not all components of y change sign. This will result in a much larger group G of invariant functions. Following papers will consider these questions.

Application is possible to linvariant problems only; here only symmetry with respect to the origin was assumed. The beginning of Section 3 showed that the existence of double inequalities (both lower and upper bound) is all that is really needed. Perhaps the method can be extended to apply to single inequalities as well.

Finally, the method derived here leads to a contraction of the action space, which is a feature it has in common with some others, like shrunken estimators and ridge estimators. It will be interesting to investigate in detail the relations between these different types of estimators.

Acknowledgement

A preliminary version of this paper was presented at the 17th European Meeting of Statisticians, Thessaloniki 1987. We are indepted to A.M.H. Gerards for his assistance in proving Lemma 2.

References

FERGUSON, T.S. (1967), Mathematical statistics, a decision theoretic approach, Academic Press, New York.

GRAYBILL, F.A. (1969), Introduction to matrices with applications in statistics, Wadsworth, Belmont California.

HERING, F., TRENKLER, G. & STAHLECKER, P. (1987) Partial minimax estimation in regression analysis, Statistica Neerlandica 41, 111-128.

JUDGE, G.G., CARTER HILL, R., GRIFFITHS, W.E., LÜTKEPOHL, H. & LEE, T.-C. (1982), Introduction to the theory and practice of econometrics, Wiley, New York.

LEHMANN, E.L. (1983), Theory of point estimation, Wiley, New York.

MOORS, J.J.A. (1981), Inadmissibility of linearly invariant estimators in truncated parameter spaces, Journal of the American Statistical Association 76, 910-915.

MOORS, J.J.A. (1985), Estimation in truncated parameter spaces, Ph. D. Thesis, Tilburg University.

TOUTENBURG, H. (1982), Prior Information in linear models, Wiley, New York.

IN 1986 REEDS VERSCHENEN

202 J.H.F. Schilderinck
Interregional Structure of the European Community. Part III

203 Antoon van den Elzen and Dolf Talman
A new strategy-adjustment process for computing a Nash equilibrium in a noncooperative more-person game

204 Jan Vingerhoets
Fabrication of copper and copper semis in developing countries. A review of evidence and opportunities

205 R. Heuts, J. van Lieshout, K. Baken
An inventory model: what is the influence of the shape of the lead time demand distribution?

206 A. van Soest, P. Kooreman
A Microeconometric Analysis of Vacation Behavior

207 F. Boekema, A. Nagelkerke
Labour Relations, Networks, Job-creation and Regional Development. A view to the consequences of technological change

208 R. Alessie, A. Kapteyn
Habit Formation and Interdependent Preferences in the Almost Ideal Demand System

209 T. Wansbeek, A. Kapteyn
Estimation of the error components model with incomplete panels

210 A.L. Hempenius
The relation between dividends and profits

211 J. Kriens, J.Th. van Lieshout
A generalisation and some properties of Markowitz' portfolio selection method

212 Jack P.C. Kleijnen and Charles R. Standridge
Experimental design and regression analysis in simulation: an FMS case study

213 T.M. Doup, A.H. van den Elzen and A.J.J. Talman
Simplicial algorithms for solving the non-linear complementarity problem on the simplotope

214 A.J.W. van de Gevel
The theory of wage differentials: a correction

215 J.P.C. Kleijnen, W. van Groenendaal
Regression analysis of factorial designs with sequential replication

216 T.E. Nijman and F.C. Palm
Consistent estimation of rational expectations models

217 P.M. Kort
The firm's investment policy under a concave adjustment cost function

218 J.P.C. Kleijnen
Decision Support Systems (DSS), en de kleren van de keizer ...

219 T.M. Doup and A.J.J. Talman
A continuous deformation algorithm on the product space of unit simplices

220 T.M. Doup and A.J.J. Talman
The 2-ray algorithm for solving equilibrium problems on the unit simplex

221 Th. van de Klundert, P. Peters
Price Inertia in a Macroeconomic Model of Monopolistic Competition

222 Christian Mulder
Testing Korteweg's rational expectations model for a small open economy

223 A.C. Meijdam, J.E.J. Plasmans
Maximum Likelihood Estimation of Econometric Models with Rational Expectations of Current Endogenous Variables

224 Arie Kapteyn, Peter Kooreman, Arthur van Soest
Non-convex budget sets, institutional constraints and imposition of concavity in a flexible household labor supply model

225 R.J. de Groof
Internationale coördinatie van economische politiek in een twee-regio-twee-sectoren model

226 Arthur van Soest, Peter Kooreman
Comment on 'Microeconometric Demand Systems with Binding Non-Negativity Constraints: The Dual Approach'

227 A.J.J. Talman and Y. Yamamoto
A globally convergent simplicial algorithm for stationary point problems on polytopes

228 Jack P.C. Kleijnen, Peter C.A. Karremans, Wim K. Oortwijn, Willem J.H. van Groenendaal
Jackknifing estimated weighted least squares

229 A.H. van den Elzen and G. van der Laan
A price adjustment for an economy with a block-diagonal pattern

230 M.H.C. Paardekooper
Jacobi-type algorithms for eigenvalues on vector- and parallel computer

231 J.P.C. Kleijnen
Analyzing simulation experiments with common random numbers

232 A.B.T.M. van Schaik, R.J. Mulder
On Superimposed Recurrent Cycles

233 M.H.C. Paardekooper
Sameh's parallel eigenvalue algorithm revisited

234 Pieter H.M. Ruys and Ton J.A. Storcken
Preferences revealed by the choice of friends

235 C.J.J. Huys en E.N. Kertzman
Effectieve belastingtarieven en kapitaalkosten

236 A.M.H. Gerards
An extension of König's theorem to graphs with no odd-K_4

237 A.M.H. Gerards and A. Schrijver
Signed Graphs - Regular Matroids - Grafts

238 Rob J.M. Alessie and Arie Kapteyn
Consumption, Savings and Demography

239 A.J. van Reeken
Begrippen rondom "kwaliteit"

240 Th.E. Nijman and F.C. Palmer
Efficiency gains due to using missing data. Procedures in regression models

241 S.C.W. Eijffinger
The determinants of the currencies within the European Monetary System

IN 1987 REEDS VERSCHENEN

242 Gerard van den Berg
Nonstationarity in job search theory

243 Annie Cuyt, Brigitte Verdonk
Block-tridiagonal linear systems and branched continued fractions

244 J.C. de Vos, W. Vervaat
Local Times of Bernoulli Walk

245 Arie Kapteyn, Peter Kooreman, Rob Willemse
Some methodological issues in the implementation of subjective poverty definitions

246 J.P.C. Kleijnen, J. Kriens, M.C.H.M. Lafleur, J.H.F. Pardoel
Sampling for Quality Inspection and Correction: AOQL Performance Criteria

247 D.B.J. Schouten
Algemene theorie van de internationale conjuncturele en strukturele afhankelijkheden

248 F.C. Bussemaker, W.H. Haemers, J.J. Seidel, E. Spence
On (v,k,λ) graphs and designs with trivial automorphism group

249 Peter M. Kort
The Influence of a Stochastic Environment on the Firm's Optimal Dynamic Investment Policy

250 R.H.J.M. Gradus
Preliminary version
The reaction of the firm on governmental policy: a game-theoretical approach

251 J.G. de Gooijer, R.M.J. Heuts
Higher order moments of bilinear time series processes with symmetrically distributed errors

252 P.H. Stevers, P.A.M. Versteijne
Evaluatie van marketing-activiteiten

253 H.P.A. Mulders, A.J. van Reeken
DATAAL - een hulpmiddel voor onderhoud van gegevensverzamelingen

254 P. Kooreman, A. Kapteyn
On the identifiability of household production functions with joint products: A comment

255 B. van Riel
Was er een profit-squeeze in de Nederlandse industrie?

256 R.P. Gilles
Economies with coalitional structures and core-like equilibrium concepts

257 P.H.M. Ruys, G. van der Laan
Computation of an industrial equilibrium

258 W.H. Haemers, A.E. Brouwer
Association schemes

259 G.J.M. van den Boom
Some modifications and applications of Rubinstein's perfect equilibrium model of bargaining

260 A.W.A. Boot, A.V. Thakor, G.F. Udell
Competition, Risk Neutrality and Loan Commitments

261 A.W.A. Boot, A.V. Thakor, G.F. Udell
Collateral and Borrower Risk

262 A. Kapteyn, I. Woittiez
Preference Interdependence and Habit Formation in Family Labor Supply

263 B. Bettonvil
A formal description of discrete event dynamic systems including perturbation analysis

264 Sylvester C.W. Eijffinger
A monthly model for the monetary policy in the Netherlands

265 F. van der Ploeg, A.J. de Zeeuw
Conflict over arms accumulation in market and command economies

266 F. van der Ploeg, A.J. de Zeeuw
Perfect equilibrium in a model of competitive arms accumulation

267 Aart de Zeeuw
Inflation and reputation: comment

268 A.J. de Zeeuw, F. van der Ploeg
Difference games and policy evaluation: a conceptual framework

269 Frederick van der Ploeg
Rationing in open economy and dynamic macroeconomics: a survey

270 G. van der Laan and A.J.J. Talman
Computing economic equilibria by variable dimension algorithms: state of the art

271 C.A.J.M. Dirven and A.J.J. Talman
A simplicial algorithm for finding equilibria in economies with linear production technologies

272 Th.E. Nijman and F.C. Palm
Consistent estimation of regression models with incompletely observed exogenous variables

273 Th.E. Nijman and F.C. Palm
Predictive accuracy gain from disaggregate sampling in arima - models

274 Raymond H.J.M. Gradus
The net present value of governmental policy: a possible way to find the Stackelberg solutions

275 Jack P.C. Kleijnen
A DSS for production planning: a case study including simulation and optimization

276 A.M.H. Gerards
A short proof of Tutte's characterization of totally unimodular matrices

277 Th. van de Klundert and F. van der Ploeg
Wage rigidity and capital mobility in an optimizing model of a small open economy

278 Peter M. Kort
The net present value in dynamic models of the firm

279 Th. van de Klundert
A Macroeconomic Two-Country Model with Price-Discriminating Monopolists

280 Arnoud Boot and Anjan V. Thakor
Dynamic equilibrium in a competitive credit market: intertemporal contracting as insurance against rationing

281 Arnoud Boot and Anjan V. Thakor
Appendix: "Dynamic equilibrium in a competitive credit market: intertemporal contracting as insurance against rationing

282 Arnoud Boot, Anjan V. Thakor and Gregory F. Udell
Credible commitments, contract enforcement problems and banks: intermediation as credibility assurance

283 Eduard Ponds
Wage bargaining and business cycles a Goodwin-Nash model

284 Prof.Dr. hab. Stefan Mynarski
The mechanism of restoring equilibrium and stability in polish market

285 P. Meulendijks
An exercise in welfare economics (II)

286 S. Jørgensen, P.M. Kort, G.J.C.Th. van Schijndel
Optimal investment, financing and dividends: a Stackelberg differential game

287 E. Nijssen, W. Reijnders
Privatisering en commercialisering; een oriëntatie ten aanzien van verzelfstandiging

288 C.B. Mulder
Inefficiency of automatically linking unemployment benefits to private sector wage rates

289 M.H.C. Paardekooper
A Quadratically convergent parallel Jacobi process for almost diagonal matrices with distinct eigenvalues

290 Pieter H.M. Ruys
Industries with private and public enterprises